This guidance was written prior to the February 27, 1997 implementation of FDA's Good Guidance Practices, GGP's. It does not create or confer rights for or on any person and does not operate to bind FDA or the public. An alternative approach may be used if such approach satisfies the requirements of the applicable statute, regulations, or both. This guidance will be updated in the next revision to include the standard elements of GGP's.

REVIEW CRITERIA FOR IN VITRO DIAGNOSTIC DEVICES THAT UTILIZE CYTOGENETIC IN SITU HYBRIDIZATION TECHNOLOGY FOR THE DETECTION OF HUMAN GENETIC MUTATIONS (GERM LINE AND SOMATIC)

Draft Version: February 15, 1996

Division of Clinical Laboratory Devices
Office of Device Evaluation
Center for Devices and Radiological Health
Food and Drug Administration

The FDA welcomes comments from the public about this document so that it can be revised as necessary to reflect generally accepted state of the art. Please send written comments to:

Peter Maxim, Ph.D.
Chief, Immunology Branch
HFZ-440
Division of Clinical Laboratory Devices
Food and Drug Administartion
2089 Gaither Road
Rockville, MD 20840-4015

TABLE OF CONTENTS

PURPOSE
DEFINITION
CLASSIFICATION
PREMARKET EVALUATION REQUIRED
DATA AND INFORMATION TO SUBMIT
I. GENERAL CONSIDERATIONS AND CLINICAL RELEVANCE
  A. Description of the clinical disorder(s) and genetic basis/mechanism(s) or other indications for use
  B. Characterization of target sequence/locus detected by the probe(s)
  C. Distribution of the target/disorder
  D. Intended use
  E. Clinical significance of test results
  F. Risk/benefit issues
  G. Clinical utility
  H. Summary of methodologies
II. DEVICE DESCRIPTION
  A. Detailed principle of the methodology
  B. Device components
  C. Manufacturing process
III. GENERAL CONSIDERATIONS FOR PMA AND 510(k) SUBMISSIONS
  A. General consideration for data collection
  B. Specific study protocol issues
    1. Sample size
    2. Sampling method
    3. Difficulties with sampling
    4. Methods of analysis
    5. Pooling of investigators' estimates
    6. Study site/investigators
    7. Assuring data integrity
IV. ANALYTIC/LABORATORY *IN VITRO* VALIDATION STUDIES FOR PMA AND 510(k) SUBMISSIONS
  A. Analytic sensitivity
  B. Specificity of probe for specific target
  C. Effects of stringency on analytic sensitivity and specificity
  D. Localization of target
  E. Interferences
  F. Precision/repeatability and reproducibility
  G. Stability data
V. DIAGNOSTIC PERFORMANCE PROPERTIES
  A. Selecting and establishing diagnostic endpoint(s)
  B. Expected values
    1. Percentage distribution of signals per cell
    2. Assay cut-off
    3. Prevalence of disorder/target

        4. Population-based location specificity
    C. Diagnostic sensitivity and specificity
    D. Methods comparison for 510(k)
    E. Predictive values
VI. REFERENCES/SUGGESTED READING

FIGURES:

Figure 1. Signal Scoring for Noncentromeric Probes in Metaphase Cells
Figure 2. Signal Scoring for Centromeric Probes in Metaphase Cells
Figure 3. Calculating the Diagnostic Performance Properties
Figure 4. ISH Probe Applications

ATTACHMENTS:

APPENDIX A: Background for *in vitro* diagnostic devices that utilize cytogenetic *in situ* hybridization technology ...

APPENDIX B: Labeling considerations for product package insert for *in vitro* diagnostic devices that utilize cytogenetic *in situ* hybridization technology ...

APPENDIX C: Glossary for review criteria for *in vitro* diagnostic devices that utilize cytogenetic *in situ* hybridization technology ...

This is a flexible document representing the current major concerns and suggestions regarding cytogenetic *in vitro* diagnostic (IVD) devices employing *in situ* hybridization (ISH) methodologies. It is based on 1) current basic science, 2) clinical experience, and 3) relevant statutes of regulation. This document will remain flexible and will be re-evaluated and revised as necessary as experience and the generally accepted state-of-the-art changes.

PURPOSE:

The purpose of this document is to assist persons who manufacture, produce, market or sponsor (hereinafter called manufacturers) cytogenetic ISH *in vitro* diagnostic devices in complying with existing labeling regulations, e.g., 21 CFR 801 and 809.10; premarket approval (PMA) requirements; and premarket notification requirements, e.g., Section 510(k) and 515 of the Federal Food, Drug, and Cosmetic Act (the Act); and other relevant statutes. This document is an adjunct to these and the FDA 87-4214 Premarket Approval (PMA) manual, and the FDA 87-4224 *In Vitro* Diagnostic Devices: Guidance for the Preparation of 510(k) Submissions Manual. It is not to supersede those documents, but is to provide additional guidance and clarification on information necessary for the FDA's premarket evaluation of these devices. The document will enable the FDA to make more informed decisions based on a uniform data base and will lead to more reliable, reproducible, and standardized commercial tests.

This document does not cover the requirements for compliance to Good Manufacturing Practices listed in publication FDA 87-4179.

DEFINITION:

The generic type device is intended for use in clinical cytogenetics laboratories as an *in vitro* diagnostic device utilizing ISH methodologies for the detection of human genetic mutations.

This document addresses devices that utilize nucleic acid probes targeted at DNA for cytogenetic ISH applications in metaphase chromosomes and intact cell nuclei (for interphase analysis). It does not address devices for *in situ* use at the tissue level in slide-mounted sections or detection of microorganisms. This document applies to ISH devices for both kit and single reagent (probe) formats.

CLASSIFICATION: Unclassified

PREMARKET EVALUATION REQUIRED:

The pathway for introducing a device into the marketplace is dependent on the level of review necessary to assure the safety and effectiveness (S & E) of the device. The type of review required [510(k) vs. PMA] will depend on the intended use of the device and special controls required to assure the S & E of the device.

Although decisions about the level of review required will be made on a case-by-case basis, in general, the FDA will entertain evaluation of those devices whose primary utility is as an adjunct to standard metaphase cytogenetic analysis for purposes of characterizing a known chromosome abnormality through the Premarket Notification [510(k)], substantial equivalence, process.

Devices whose primary utility or intended use is as a "stand alone" based on interphase analysis ("dot counting") for test reporting will generally require a Premarket Approval (PMA). Such applications could impact significantly on changing the current standard of practice for cytogenetic analysis and thus raise new issues of safety and effectiveness. Examples of products that may fall into this category include centromeric probes for chromosomes commonly associated with constitutional aneuploidies (e.g., chromosomes 13, 18, 21, X, and Y) and recurrent aberrations associated with certain malignancies.

In making a substantial equivalence (SE) determination, the Safe Medical Devices Act of 1990 (SMDA '90) requires the comparison of a new device to a legally marketed predicate device. For devices that utilize new technology, the device must be as safe and effective as the predicate device and there must be no new issues of safety and effectiveness raised by the device's new technology.

DATA AND INFORMATION TO SUBMIT:

As with any evolving technology, numerous issues emerge in the review of molecular diagnostic methods. It is incumbent upon the manufacturers to provide evidence that the device is safe and effective for its intended use/indications for use with clinical utility. Appropriate information and data are required to validate and characterize the performance properties of the device for the stated intended use. Statements and claims made throughout the submission should be supported with key literature citations and/or validated with relevant data.

In addition, all information relevant to performing the assay and interpreting test results should be included in the package insert (PI). For further information regarding labeling requirements, see Appendix B.

When applicable, manufacturers should address the issues and include the general information discussed below in their submissions:

I. GENERAL CONSIDERATIONS AND CLINICAL RELEVANCE

Provide the following information about the device:

A. DESCRIPTION OF THE CLINICAL DISORDER(S) AND GENETIC BASIS/MECHANISM(S) OR OTHER INDICATIONS FOR USE

1. Describe any known recurring aberrations or clinical disorders (e.g., phenotypic, chromosomal, or genotypic characteristics) associated with the target sequence being detected with the device.

2. Describe the genetic basis and mechanisms that are relevant to the expression of the clinical disorder including inheritance pattern as validated by family studies, e.g., germ line vs. somatic; mendelian vs. chromosomal; dominant vs. recessive; autosomal vs. X-linked; numeric vs. structural chromosome abnormality; if possible, and any other relevant genetic mechanism pertinent to expression of the trait, e.g., genetic heterogeneity, reduced penetrance, variable expressivity, delayed onset, uniparental disomy, imprinting, mosaicism, etc.

3. Describe any known genetic heterogeneity for the clinical disorder of interest.

B. CHARACTERIZE THE TARGET SEQUENCE/LOCUS DETECTED BY THE PROBE(S) TO INCLUDE THE FOLLOWING INFORMATION

Each target locus/allele/mutation being detected by the test should be fully characterized. Characterization should include the following, as applicable:

1. State what, specifically, the probe detects, e.g., the molecular defect associated with the disorder of interest; whether a chromosome rearrangement site, deletion, duplication, inversion, expansion/amplification, etc. is being detected.

2. Provide evidence that the target locus is well documented in the scientific literature or appropriate data bases. Identify the genomic map position of the locus using standard nomenclature from the Human Gene Mapping Workshop Nomenclature Committee, Geneatlas, Genome Data Base (GDB), and/or the International System for Human Cytogenetic Nomenclature (ISCN), 1985 and Cancer Supplement, 1991, etc. Indicate whether multiple probes are recommended for detecting the aberration of interest.

3. Give the rationale for selecting the sequence chosen to be the target for detection. Clarify whether the probe is targeted directly at the

mutant DNA sequence being identified or whether it is directed at a flanking region. If detection is directed at a flanking region, provide relevant information about the physical map and demonstrate that the probe is directed at the appropriate target to detect the mutation pertinent to the particular clinical or genetic state of interest.

4. Provide information to substantiate application of ISH to detect the target sequence/disorder in question; cite appropriate literature references.

5. For probes used to detect microdeletions, provide evidence that the probe is directed at the smallest known region of overlap (SRO) in patients that exhibit the associated clinical phenotype. Such evidence is a function of the number of patients with the deletion in question that are tested during the initial validation studies generated with the device or reported in the literature. Since there may be rare individuals who have even smaller regions or in situations in which the smallest region is not completely understood, describe the "detection limits" of the probe and include this information in the package insert.

6. State the size (number of nucleotides) and base sequence of the target, if known.

7. Discuss the polymorphic nature of the target sequence(s)/loci(us).

8. Discuss the degree of sequence/locus conservation if relevant.

C. DISTRIBUTION OF THE TARGET/DISORDER

1. Discuss whether the target being detected is present in normal, carrier, or only in phenotypically affected individuals.

2. Describe the population(s) at risk for the target/disorder in question.

3. Describe the distribution of the target/disorder in the population(s) of interest, frequency of the defect, mutation rates, and any known variability in the frequency distribution between demographically defined subpopulations, e.g., geographic, racial, ethnic, etc., or explain why this information is not relevant or available.

D. INTENDED USE

Describe the intended use of the device and provide the following information:

1. State whether the test will be used as an adjunct to metaphase analysis or as a stand alone for interphase ISH test reporting.

2. Define the target populations to be tested with the device and the test setting.

Describe how the test will be used in the target populations, e.g., whether for screening, diagnosis, prognosis, detecting minimum residual disease, monitoring therapy, etc.

Describe (for FDA) all known potential uses for which the device could be used, even if FDA clearance/approval is not requested for that use and state the key restrictions, limitations, and contraindications for use of the device in the limitations section of the PI.

3. Types of acceptable test specimens (matrix)

List all specimen types/matrix(ices), tissue source, or special preparations (e.g., cultured specimen) required for use with the methodology(ies) of this device. Discuss problems incurred with the use of any listed specimen types that will significantly impact on test results and interpretation.

For example, if prenatal use is intended, address the consequences of an increased incidence of mosaicism in chorionic villi samples as they relate to the reliability of the device. State the exact type of specimen required (e.g., amniotic fluid, chorionic villi, fetal cells from maternal blood, etc.), and the optimal gestational age for testing.

E. CLINICAL SIGNIFICANCE OF TEST RESULTS

Discuss the clinical implications and significance of test results. Discuss the significance of test results in terms of the probability that the individual will be clinically affected and whether the test diagnosis will have an effect on health or life.

F. RISK/BENEFIT ISSUES (for all novel, unproven, and new intended use applications)

Assess the probability for and the significance of false positive/negative test results vs. benefit of information provided. Also, consider how erroneous results will impact on patient management and care, family members, for prenatal vs. postnatal, or constitutional vs. acquired disorders, etc.

Additional variables to assess include prevalence of the disease; burden of the disease; acceptance of testing; speed of testing; predictive value of a test, e.g., screening vs. diagnosis; etc.

In general, a new device should be introduced for diagnostic use only if the performance properties are demonstrated to be as good as existing methods or if the new method provides a significant new advantage, e.g., increased accessibility, improved turn around time, etc.

G. CLINICAL UTILITY

Establish that the device has clinical utility for each stated intended use and indications for use. The clinical utility of the intended use should be recognized by professionals in the field, supported by peer reviewed literature, and/or data generated by the manufacturer about the device to demonstrate that use of the device will yield clinically relevant information, i.e., that the test will contribute significantly to the diagnosis of a disease or the genetic status of the individual or family member tested.

H. SUMMARY OF METHODOLOGIES

Provide a historical summary of all diagnostic methodologies used to detect the disorder in question. Describe the merits/advantages and limitations/disadvantages of the device methodology(ies) within the context of other available methodologies.

II. DEVICE DESCRIPTION

Submit all relevant descriptive information about the device necessary for an effective evaluation of the device. Additionally, include all information relevant to performing the assay and for interpreting test results in the appropriate section of the accompanying PI.

A. DETAILED PRINCIPLE OF TEST METHODOLOGY

Describe how the device functions, the basic scientific concepts that form the basis for the device, and the significant physical and operative characteristics.

Provide a thorough explanation of all aspects of the test methodology.

1. Describe the biological/chemical principle of the hybridization reaction. Discuss factors that affect the hybridization reaction(s), which are unique to the product, when applicable.

2. Describe the molecular interactions and methodology used for probe detection.

B. DEVICE COMPONENTS

Characterize the functional components or ingredients of the device and include pertinent information in the Reagents section of the PI. Note: All proprietary information submitted should be clearly designated as confidential/proprietary.

1. Detection probe:

Describe the following aspects of all detection probes used in the device: probe type (DNA/RNA); size by appropriate methods; base sequence, if pertinent; restriction enzyme map, if pertinent; probe label - isotopic (e.g., ^{32}P) or non-isotopic (e.g., biotin); manner of attachment of detection molecules; stability of attachment; function of multiple probes, if applicable; etc.

Demonstrate probe purity or characterize any impurities present and state why they will not interfere with the assay.

2. Components of the probe detection system

Provide detailed characterization of all components used to detect hybridization products, e.g., reporter molecule, chromophors, enzymes, antibodies, conjugates, substrates, etc.

3. Control material

When feasible, manufacturers are encouraged to provide validated controls with the device that would facilitate widespread use of the comparable standards and aid inter-laboratory precision/reproducibility and accuracy. Controls should cover the medically relevant range of reportable values and especially stress the system near the cutoff and limits of detection (e.g., minimal percentage of abnormal cells detectable).

Control materials should be derived from an appropriate specimen type, e.g., the same specimen type as recommended for use in the labeling may be desirable.

Characterize the performance properties of any control materials provided.

4. The manufacturer may wish to consider including internal controls to tag the chromosome/region of interest and to monitor the efficiency of hybridization of the system, when appropriate.

5. Characterize additional reagents, components, or equipment provided or recommended for use with the device, and describe their function in the assay.

6. Sequence accession number if registered with data bank and any relevant licensure/patent information.

C. MANUFACTURING PROCESS

Provide a description of reagent specifications, manufacturing process, and quality control for relevant components of the device, e.g., describe probe production: cloning vector, vector size and verification of size, percent insert, isolation, purity, conjugate/attachment, cloning enzyme, resulting polymorphism, optimal concentration, functional validation, and how the probe is generated as is relevant or practically possible.

Standard Operating Procedures (SOPs) must be in place for manufacturing production lot components used to generate data to support investigational studies of 510(k)s and PMAs.

III. GENERAL CONSIDERATIONS FOR PMA AND 510(k) SUBMISSIONS

Device validation must be conducted before a new device is introduced for *in vitro* diagnostic (IVD) use. The test should be subjected to (1) thorough literature review, (2) analytic studies, and (3) laboratory and clinical/diagnostic correlation studies to: (a) characterize the locus/mutation(s) being detected, (b) establish the performance properties of the device to insure the device's ability to provide consistent and reliable results, (c) establish the clinical utility of the test, (d) define aspects of the procedure which must be carefully regulated to maintain device performance, and (e) define certain limitations of the test. Such validation is necessary to assure the safe and effective use of a device for its intended use.

The FDA requests different types and amounts of data and statistical analyses in applications to market *in vitro* diagnostic devices. The amount and type of data requested depends on: 1) the test analyte/target of interest (e.g., the potential clinical or diagnostic relevance of the abnormality detected by the device), 2) intended use which influences whether the application is a 510(k), or an original Premarket Approval Application (PMA), 3) whether the test is qualitative or quantitative, 4) the study design, and 5) certain claims made by the manufacturer.

Performance claims made in the product labeling must be substantiated with submission of supporting literature and/or appropriate studies. Generally, novel and unproven applications require more stringent validation that includes characterization of both the analytic and diagnostic performance properties. In general, diagnostic validation will be required for devices whose primary stated and/or implied utility is for disease/mutation specific detection. Devices with a primary utility for characterizing non-recurring aberrations such as probes directed at telomeric sequences, whole chromosome paints, etc., will generally require only characterization of the analytic performance properties described in Section IV and the population based target localization data in Section V.B.4.

All performance claims should be documented by studies that set quantitative goals for performance of the device and consider all important sources of variability, using statistical procedures that quantify their effects. Data are statistically sufficient if there are enough data gathered under the right conditions to give a high degree of confidence that the procedure will work in appropriate laboratory settings and intended uses of the device. The data and statistical analysis should be adequate to demonstrate the safety and effectiveness of the device for all claimed intended uses/indications for use. Detailed protocols for device validation are beyond the scope of this document. This section will provide guidance of key points to consider for validation purposes.

A. GENERAL CONSIDERATIONS FOR DATA COLLECTION:

1. Generate data to support claims for each intended cell stage (e.g., metaphase vs. interphase), specimen type (e.g., blood, chorionic villus (CVS), fibroblasts, etc.), and specimen preparation method (e.g., cultured vs. uncultured), as relevant, or provide evidence that these variables do not effect the performance properties between specimen types or preparation methods.

2. Process and analyze cells from an appropriate source consistent with the intended use. For example, unstimulated cells from appropriate (affected) tissues may be suitable for analysis in certain malignant disorders. In other cases, certain growth factors and mitogens may improve the user's ability to identify abnormal clones (e.g., B-cell disorders).

3. Perform testing to establish the performance characteristics of the device at external user sites according to the intended use and instructions provided in the product labeling.

4. All studies must be performed in compliance with Institutional Review Board requirements (with informed consent, when relevant) and Investigational Device Exemption regulations, when applicable.

5. Generate performance data for both preclinical and clinical studies using a test device that is standardized in its composition or design and performance [21 CFR 860.7(f)(2)].

6. Provide detailed protocols for each study used to establish the relevant analytic and clinical/diagnostic performance properties of the device discussed in sections IV and V. This includes protocols (or literature citation) of any reference methods/devices to which the device is compared. The details described in the protocols must be followed consistently throughout the studies and between sites; any deviations must be addressed. For example, consistent criteria for ISH cell selection, scoring, analysis, and interpretation should be used throughout.

7. Establish strict cell selection and acceptability criteria. For example, what is a "readable" cell? Generally, all metaphase cells have at least one normal chromosome of a pair (for autosomes) and therefore should have one chromosome with a positive signal. If a normal chromosome is not present, the metaphase may not be suitable for scoring. Also, establish the criteria for dealing with uninterpretable results, e.g., due to weak signals.

8. Perform all scoring of signals, masked, when possible, i.e., the person scoring is different from the one who performs the procedure and they have no knowledge of the assay conditions or patient information, e.g., diagnostic status. Provide details of the measures taken to assure that all studies and analyses were performed in a masked fashion.

9. Establish and clearly define the criteria for scoring ISH signals by visual, microscopic examination of interphase/metaphase cells.

When scoring chromatid signals in metaphase cells, distinguish between: absence of signal and detection of signal on one or both chromosomes of a pair for centromeric probes and between absence of signal and detection of signal on one or both chromatids of each chromosome of a pair as appropriate for each cell analyzed/examined. For example, when scoring metaphase probes that do not include an internal control, a positive signal may be defined as a signal on both chromatids of the chromosome or spanning the centromere for an alpha satellite probe; a signal on one chromatid would not be scored as a positive signal (see Figures 1 and 2).

For interphase cells, strict cell inclusion/exclusion and scoring criteria should be established to reduce the variability of interpretation and increase the reliability of the test. For example, it may be appropriate to examine and score the proportion of cells with 0, 1, 2, 3, 4, etc., signals and the proportion of any uninterpretable/unreadable cells. In many cases, scoring cells with no signal may be counter productive since nullisomies are unlikely. When appropriate, categories may be individualized depending on what degree of numerical abnormality is expected (diploid, trisomic, triploid, tetraploid, etc.), how many probes are combined, and whether multicolor signals are being scored. In other cases, it may be appropriate to categorize scoring of greater than expected, equal to expected, less than expected, and unreadable.

10. Provide the rational for selecting the number of cells selected for analysis for the studies submitted. The minimum number of chromosome spreads and nuclei per specimen required to establish the performance properties of the device depends on the desired precision of the estimate (e.g., percent of tri-signaled nuclei), the accuracy of classification (e.g., errors in determination of relevant event), and the probability of locating rare signals (e.g., minimal residual disease). In addition, the intended use (e.g., whether the product is used to evaluate for mosaicism, residual disease, etc.), the availability of metaphase spreads, the test's ability to distinguish one state of nature from another, and the type of specimen may be relevant. The number of observations required for the manufacturer to characterize the device should be sufficient to provide precise estimates of the performance properties of the device. This could differ substantially from what is required by the user for generating individual analyses.

User requirements will depend on the analytic performance of the probe (e.g., efficiency of hybridization) as determined by the manufacturer and verified by the user. For example, based on a probe's hybridization efficiency, the manufacturer may make the statement that if a given number of random cells are tested then, with a stated confidence (e.g., 95 percent), mosaicism of a given percent will be detected.

The efficiency of hybridization is dependent not only on the probe but also on the target material. Variability in the quality of cytogenetic preparations may impact on the results obtained with the probe just as it impacts on the quality of G-banding. It may be useful for the manufacturer to establish reference tables that will allow the user to pose the following type of question: "Given that the hybridization efficiency is X percent, how many cells must be analyzed to have some stated probability Y of detecting an abnormality of interest, e.g., mosaicism at a level of Z percent?" The manufacturer could use the same reference tables to characterize subsequent probes.

11. Present a summary of data with statistical analyses and conclusions to permit independent analysis by the FDA. Address the statistical treatment of each category of signal scores as described in II.A.8. above, including unreadable/uninterpretable cells. When appropriate, charts [scattergrams, histograms, receiver operating characteristics (ROC) curves, etc.] may be used as part of the analyses and conclusions.

Summarize data from interphase and metaphase photographs in tabular, diagrammatic, or narrative form. Ideally, photographs of examples of expected results should be submitted; Xerox copies of FISH are generally uninterpretable.

12. Include, in the PI, a summary of the study protocol(s), data analysis, and conclusions for each relevant performance property discussed in Sections IV and V below.

B. SPECIFIC STUDY PROTOCOL ISSUES

The purpose of clinical studies is to demonstrate the safety and effectiveness of the device for its intended use (medical and diagnostic claims) when accompanied by adequate directions and warnings against unsafe use. *In order for a premarket applications to succeed, appropriate and detailed planning of the protocol is absolutely essential.* The manufacturer is encouraged to contact the FDA's reviewing division to discuss the study protocol before initiating clinical studies and to consult with a biostatistician during the initial planning stages of the study and during data analysis.

This section addresses points to consider when designing study protocols. This is intended to supplement, not replace, the FDA 87-4214 Premarket Approval (PMA) manual.

For the studies described in section IV. and V. below, provide the following additional information, as applicable:

1. Sample size:

The minimum number of subjects required depends on what is to be proven. It will vary with the specific target sequence being detected, indications for use, availability of specimens, error rate, precision of estimate required, etc.

Provide a complete description of the sample size determination. Ideally, the sample size will be based on a null hypothesis of major importance to the study, an effect or departure from null hypothesis conditions that would be clinically meaningful, and a choice of significance level (alpha, typically 5%) and desired level of probability of detecting such an effect if it exists (i.e., "power" or 1 - beta). Although a power of 80% is frequently used in [non-genetic] statistical analyses, a greater level of statistical power (e.g., 95 percent or higher) may be required when failure to reject the null hypothesis has serious consequences or when the purpose of statistical analysis is to establish equivalence.

Prior to beginning clinical/diagnostic studies, plan the sample size that will be large enough to establish the clinical/diagnostic performance properties, e.g., sensitivity and specificity, of the device with sufficient precision so that predictive values together with their confidence limits can be adequately estimated for a range of population prevalence. The level of confidence selected should be justified taking into consideration the consequences of wrong decisions that might be made due to incorrect test results. For example, if the consequences of misdiagnosis are very grave, then one would like to have at least 99 percent confidence that the corresponding predictive value is greater than some quantity. On the other hand, if the consequence is less grave, then one might be satisfied with knowing, with only 90 percent confidence, that the predictive value is some specific quantity. Otherwise, the 95 percent confidence interval may be required.

In the hypothesis testing situation, the sample size is based on the number of subjects needed to achieve a predetermined, minimally detectable difference that is clinically meaningful for each hypothesis tested. Consider effect criteria, alpha (type I) and beta (type II) statistical error tolerances, anticipated variance of the response variables, and any assumptions or statistical formulas that are required for sample size determination.

Also, determine the number of subjects to be selected at each study site. The samples size should be large enough at each site to permit separate analysis and conclusions (see pooling of data in III.B.5. below).

2. Sampling method:

Specify the type of statistical sampling plan to be used in selecting specimens for analysis from subjects in the target population. Identify characteristics of the target population that may impact on the method of sampling needed and the particular method of analysis. The makeup of the target population with regard to prevalence, ethnicity, relatives vs. non-relatives, etc., should be considered. For example, it would be necessary, should relatives be used to determine the diagnostic sensitivity and specificity, to use a method of statistical analysis that allows for the fact that there are familial correlations.

Appropriate subject sampling criteria and statistical methods must be carefully developed to avoid introducing biases, e.g., subject inclusion/exclusion criteria, masking (e.g., different personnel performing test from those selecting samples), etc. For example, estimates of diagnostic sensitivity and specificity are typically based on the independence assumption of all study specimens/subjects included for analysis. Therefore, inclusion of multiple specimens per subjects may skew results toward achieving artificially high associations and artificially low standard deviations. Avoid biases introduced by inclusion of multiple specimen(s) types per subject or design an analytical method for removing bias in the key endpoint statistics, e.g., bootstrapping or other resampling procedure.

Prospective studies to support the intended use/indications for use of the test are desirable although studies performed with archived specimens (retrospective studies) may provide appropriate information. Describe measures taken to avoid all potential biases introduced by retrospective sampling instead of prospective sampling.

In addition, if retrospective studies are used, the manufacturers and investigators should be cognizant of the legal/ethical issues that may arise from use of such specimens and the potential obligation to inform the subject if the new test provides information that may potentially have a fundamental impact on that individual if discovered. In general, if personal identifiers are linked to the specimen, the investigator/manufacturers may have the duty to inform regardless of whether the specimen was initially obtained and studied for the same or different purpose.

Use an appropriate (probabilistic) sampling plan to ensure the representativeness and unbiasedness of the sample data to the target patient population regardless of whether the sampling is prospective or retrospective. For example, if the target population includes subjects with ambiguous cytogenetic results, it is not appropriate to exclude them from the study sample.

The inclusion/exclusion criteria should be robust enough to avoid enlisting a disproportionately large number of subjects who will ultimately be disqualified. Exclusions of eligible subjects from the study or any inclusions of ineligible subjects into the study cannot be justified. There must be a full accounting for all subjects entered into the study.

3. Difficulties with sampling:

The FDA recognizes the difficulty encountered by manufacturers in obtaining a sufficient number of subjects for studies to validate devices intended for use in detecting very rare disorders. In cases of low prevalence disorders or disorders with a long lag-time between testing and onset of disease, it may not be feasible to test the diagnostic performance of the tests in the intended use population. It has been recommended that in these special situations, a reasonable approximation of the tests performance may be determined using subjects with overt disease to test for sensitivity and non-affected subjects beyond the age at which the disease usually becomes manifest to estimate specificity (Andrews, et al.). Test interpretation must take into account any biases in estimates introduced by less than ideal study design. Also, explain any difficulty in obtaining adequate numbers of informative specimens.

4. Methods of analysis:

Provide details for the statistical methods and analyses, and corresponding computer outputs and references so that independent verification of the findings can be made. All statistical analysis should be based appropriately on the mode of sampling used and the variables being measured.

Describe and justify the assumptions, results, and the statistical methods used for the type of data submitted, e.g., qualitative discrete data, quantitative continuous data, etc., and the distribution of data (paired vs.

independent)

Traditionally, diagnostic sensitivity and specificity are based on binary (dichotomous) outcomes: presence or absence of disease and positivity or negativity of test results. For genetic disease testing, a device may be used to delineate multiple (>2) "states of nature" relevant to the target sequence of interest. Therefore, dichotomous outcomes often are not applicable; but rather, polytomous (n>2) ones are applicable. In such cases, it may be appropriate to calculate the probability that the device detects the state of interest, given that the subject has the particular state. The states of nature relevant to the submitted device must be clearly stated and defined. Present test data with estimated parameters and their confidence intervals, any relevant estimates of error, and conclusions. Describe the statistical method(s) used to determine them. All categories of test results, e.g., equivocal or uninterpretable, positive, negative, etc., must be accounted for in the data analysis.

For submissions with diagnostic claims, provide the working data and, if possible, copies of raw data as a ASCII file on a DOS-formatted diskette. Provide a complete key for interpreting any data and information submitted.

5. Pooling of investigators' estimates:

Present test data with analysis and conclusions for each site/investigator. Estimates can be pooled over investigators, if pooling is statistically and clinically justified. The homogeneity assumption of test results over all relevant subgroups should be statistically validated before pooling estimates among sites. Estimates should be pooled by taking a weighted average of the site-specific estimates. The weights used to calculate that average should reflect the precision of the site specific estimates. If pooling of estimates is not justified, more data rather than less data will be desirable and the results should be presented site-by-site. Data from various sites may not generally be pooled and analyzed as if they came from a single site.

6. Study sites/investigators:

a. Select study sites to represent the spectrum of demographic features, "disease" prevalence, geographical locations, etc. Testing sites should be representative of the setting indicated for actual testing. For PMAs, use at least three independent investigators at separate sites; at least one study should be conducted in the United States.

b. Investigations should be performed by investigators unaffiliated with the manufacturer. The manufacturer should assure that appropriate steps were taken to assure that individuals performing tests for the studies have appropriate experience/training.

Identify test sites by institutional name and location; include the name, title, and phone number of the responsible investigator at each site.

c. Plan the duration of both analytical and clinical studies and provide the beginning and ending dates for clinical/diagnostic studies.

7. Assuring data integrity:

It is the manufacturer's responsibility to oversee all aspects of the evaluation studies to assure the integrity of the data for both in-house and external evaluation site studies. Suggestions are:

a. to appoint a study coordinator to oversee all aspects of the clinical studies. This individual must have knowledge of all details of the study and serves to protect the integrity of the data.

b. to provide a detailed, written protocol applicable to all study sites before subject recruitment begins. It is the responsibility of the study coordinator/manufacturer to assure that all study sites adhere to the study protocol(s).

c. to justify, document, and address any and all deviations from the protocol and defined statistical sampling methods once the study begins.

IV. ANALYTIC/LABORATORY *IN VITRO* VALIDATION STUDIES FOR PMA AND 510(k) SUBMISSIONS

Identify and characterize the critical analytic aspects of the procedure that must be carefully controlled and monitored to provide consistent and reliable results. Perform studies and provide data to demonstrate the analytic performance of the device for the properties listed below, as applicable. Validate the procedure for all specimen categories that will be utilized for testing. If multiple tissues/specimen types are acceptable for testing, document that different tissues from the same individual will yield the same results. Actual patient specimens should be used when possible. In some cases, e.g., when patient specimens are not readily available, simulated specimens may be acceptable for use. The specimen type for both simulated and patient specimens should be the same as indicated in the intended use.

A. ANALYTIC SENSITIVITY

1. Definition:

Analytic sensitivity for ISH applications may be defined as the ability of the device to detect a particular target, e.g., chromosome, locus, sequence, of interest. This may be quantitated by determining the proportion of available targets that are detected by the device, i.e., the probability (P) that the target sequence (TS) is detected, given that it is present, i.e., P(TS is detected | TS is present)

Another measure of sensitivity that is relevant for most ISH applications is the level of mosaicism or minimum residual disease that can reliably be detected with the device, i.e., the minimum percentage of "positive" cells in a specimen that can reliably be detected.

2. Studies:

Provide documentation that the specimens used for analysis have the target of interest. This could be demonstrated in metaphase cells using double analysis (ISH followed by banding). ISH analysis will determine how well the probe hybridizes to the target; banding studies will show whether the relevant chromosome band in which the target sequence is located is present. It may be appropriate in some cases to further characterize the specificity of the probe by subjecting genomic DNA from the tissue under study to Southern analysis.

Provide data to demonstrate intraspecimen signal detection frequencies using appropriate cell types under recommended stringency conditions. When determining the true proportion of chromatids/chromosomes/cells that contain the correct number of target sequences, one must consider the disease state of the individual, the tissue type of the specimen, cell type, and the stage of the cell cycle in which the studies are carried out. The proportion of cells of interest may vary between different stages of the cell cycle (metaphase vs. interphase) depending on the proliferative advantage/disadvantage of certain cell types within a specimen.

In order to arrive at a reliable estimate of sensitivity, study specimens from a sufficient number of individuals known to have (or to have lost) the target/mutation of interest.

3. Score signals as stated in III.A.8. above.

4. Report, as appropriate, results for metaphase cells as indicated below and give confidence intervals for each estimate:

a. Average number of signals per chromatid

$$\frac{\text{Total \# signals in target chromatid bands}}{\text{Total number of target chromatids available}}$$

b. Average number of signals per chromosome

$$\frac{\text{Total \# signals in target chromosome bands}}{\text{Total number of target chromosomes available}}$$

c. Average number of signals per cell

$$\frac{\text{Total signals in target chromosome bands}}{\text{Total number of cells examined}}$$

B. SPECIFICITY OF PROBE FOR SPECIFIC TARGET

1. Definition: Analytic specificity refers to the ability of the device to distinguish the target from other sequences in the specimen. The concept of specificity as it relates to ISH can be evaluated at several levels.

Parameters of interest include: 1) measurement of the proportion of cell with false positive signals due to cross-hybridization and 2) characterization of the cross-hybridization.

Another measurement of interest may be the probability that no signal is detected, given that the target sequence (TS) of interest is not present:

$$P(\text{no signal detected} \mid \text{TS not present})$$

2. Demonstrate in metaphase spreads (e.g., using double analysis, ISH followed by banding) that the probe(s) hybridize(s) to the claimed chromosome/locus/region and that there is no significant hybridization to non-target under recommended stringency conditions. For example,

consider whether the cross-hybridization is random vs. non-random, what degree of homology exists between target and hybridizing non-target, etc.

3. Score results as stated in III.A.8. above for each cell and specimen studied.

4. Report, as appropriate, the results for metaphase cells as indicated below:

   a. proportion of cells in a specimen that demonstrate signal detection at sites other than the target;

   b. the number of different hybridizing non-target chromosomes/regions/loci per cell/specimen, whether random or non-random; and

   c. using standard nomenclature, identify the chromosome/chromosomal location of all hybridizing non-target and, when possible, characterize these non-target sequences.

5. State acceptability criteria for analytic specificity. For example, what level of hybridization to non-target is permissible, e.g., present in $\leq$ 2 percent of cells for both whole chromosome probes and repeated sequence probes (Laboratory Standards and Practices Guidelines, 1995, American College of Medical Genetics)?

C. EFFECT OF STRINGENCY ON SENSITIVITY AND SPECIFICITY

Sensitivity and specificity are determined by the degree of binding between the target and probe nucleic acids which in turn may be affected by the assay stringency and the degree of homology to other sequences. Sensitivity and specificity may be optimized with appropriate conditions of stringency. The degree of stringency is regulated by varying conditions of the hybridization reaction. Varied [non-optimal] conditions of stringency may be used to determine the tolerance of the assay for permitting cross-hybridization to non-target sequences. This can best be evaluated utilizing metaphase cells.

Provide data to demonstrate intraspecimen signal detection frequencies of the probe under varying stringency conditions, e.g., temperature, salt, etc. Demonstrate reactivity with specimens from individuals with known genotype. Perform studies and report results as described in III.B.1. and III.B.2. above.